

The Steubenville Comprehensive Air Monitoring Program (SCAMP): Overview and Statistical Considerations

Daniel P. Connell, Jeffrey A. Withum, Stephen E. Winter, and Robert M. Statnick

CONSOL Energy Inc. Research & Development, South Park, PA

Richard A. Bilonick

CONSOL Energy Inc., Pittsburgh, PA

ABSTRACT

Average concentrations of particulate matter with an aerodynamic diameter less than or equal to $2.5\ \mu\text{m}$ ($\text{PM}_{2.5}$) in Steubenville, OH, have decreased by more than $10\ \mu\text{g}/\text{m}^3$ since the landmark Harvard Six Cities Study¹ associated the city's elevated $\text{PM}_{2.5}$ concentrations with adverse health effects in the 1980s. Given the promulgation of a new National Ambient Air Quality Standard (NAAQS) for $\text{PM}_{2.5}$ in 1997, a current assessment of $\text{PM}_{2.5}$ in the Steubenville region is warranted. The Steubenville Comprehensive Air Monitoring Program (SCAMP) was conducted from 2000 through 2002 to provide such an assessment. The program included both an outdoor ambient air monitoring component and an indoor and personal air sampling component. This paper, which is the first in a series of four that will present results from the outdoor portion of SCAMP, provides an overview of the outdoor ambient air monitoring program and addresses statistical issues, most notably autocorrelation, that have been overlooked by many $\text{PM}_{2.5}$ data analyses. The average $\text{PM}_{2.5}$ concentration measured in Steubenville during SCAMP ($18.4\ \mu\text{g}/\text{m}^3$) was $3.4\ \mu\text{g}/\text{m}^3$ above the annual $\text{PM}_{2.5}$ NAAQS. On average, sulfate and organic material

accounted for ~31% and 25%, respectively, of the total $\text{PM}_{2.5}$ mass. Local sources contributed an estimated $4.6\ \mu\text{g}/\text{m}^3$ to Steubenville's mean $\text{PM}_{2.5}$ concentration. $\text{PM}_{2.5}$ and each of its major ionic components were significantly correlated in space across all pairs of monitoring sites in the region, suggesting the influence of meteorology and long-range transport on regional $\text{PM}_{2.5}$ concentrations. Statistically significant autocorrelation was observed among time series of $\text{PM}_{2.5}$ and component data collected at daily and 1-in-4-day frequencies during SCAMP. Results of spatial analyses that accounted for autocorrelation were generally consistent with findings from previous studies that did not consider autocorrelation; however, these analyses also indicated that failure to account for autocorrelation can lead to incorrect conclusions about statistical significance.

INTRODUCTION

Concerns about airborne particulate matter (PM) with an aerodynamic diameter less than or equal to $2.5\ \mu\text{m}$ ($\text{PM}_{2.5}$) have arisen largely from the results of epidemiological studies that suggest a positive association between ambient average $\text{PM}_{2.5}$ concentrations and adverse health effects. Particulate air pollution has been linked to mortality^{1,2} and to cardiovascular and respiratory morbidity.³⁻⁵ In response to these findings, the U.S. Environmental Protection Agency (EPA) announced a new National Ambient Air Quality Standard (NAAQS) for $\text{PM}_{2.5}$ in July 1997, which set an annual limit of $15\ \mu\text{g}/\text{m}^3$ (averaged over 3 yr) and a 24-hr limit of $65\ \mu\text{g}/\text{m}^3$ (98th percentile, averaged over 3 yr).⁶

Steubenville, OH, has long been a focus of particulate air pollution studies. Located along the Ohio River ~36 miles (58 km) to the west of Pittsburgh, Steubenville is situated in a highly industrialized area that includes coke plants, metal smelting and processing plants, and coal-fired electric power plants. Rural areas surround the city, and the region experiences distinct seasons characterized

IMPLICATIONS

Steubenville, OH, is a key research location for $\text{PM}_{2.5}$. The city was a focus of epidemiological studies that discovered associations between PM and adverse health effects, leading to the promulgation of an NAAQS for $\text{PM}_{2.5}$. However, its population, industrial activity, and air pollution levels have declined appreciably since these studies were conducted. Nevertheless, Steubenville will likely not attain the annual $\text{PM}_{2.5}$ NAAQS. The Steubenville Comprehensive Air Monitoring Program provides a much-needed current and statistically sound assessment of $\text{PM}_{2.5}$ in the Steubenville region. Analysis of its data will aid in the development of $\text{PM}_{2.5}$ reduction strategies for Steubenville and similar areas as well as provide insights relevant to the design and interpretation of $\text{PM}_{2.5}$ health effects studies.

by cold winters and mild summers. Hence, the location is ideal for examining the effects of season and urban versus rural location on air pollutant concentrations. Historically, levels of $PM_{2.5}$ at Steubenville have been higher than the standards promulgated by the EPA in 1997. Spengler and Thurston⁷ measured a mean $PM_{2.5}$ concentration of $36.1 \mu\text{g}/\text{m}^3$ between April 1979 and July 1981, and Laden et al.⁸ reported a mean $PM_{2.5}$ concentration of $30.5 \mu\text{g}/\text{m}^3$ between 1979 and 1988.

Some important epidemiological studies have focused on the Steubenville region. The landmark Harvard Six Cities Study¹ reported an association between Steubenville's elevated ambient $PM_{2.5}$ concentration and increased mortality. Schwartz and Dockery⁹ reported a positive correlation between total suspended particulates and daily mortality at Steubenville, and Schwartz et al.¹⁰ found a statistically insignificant association between $PM_{2.5}$ and daily mortality.

However, Steubenville has changed appreciably since these studies were conducted. The Steubenville-Weirton Metropolitan Statistical Area (MSA) lost 4,200 manufacturing jobs during the 1990s, largely due to the decline of the steel industry,¹¹ and experienced the largest percentage decline (7.4%) in population among MSAs in the United States.¹² Data from the early 1970s through 1990 suggest a 5% per year decline in $PM_{2.5}$ concentrations in Steubenville,¹³ indicating that air pollution in the city also has changed substantially. Nevertheless, Steubenville continues to be an industrial center. Although the Steubenville-Weirton MSA (population 132,000) accounts for less than 0.05% of the total U.S. population, it accounted for more than 1% of total U.S. emissions of CO, NO_x , and SO_2 from major industrial facilities in 1999, according to estimates from the EPA's National Emission Inventory Database.¹⁴ Hence, given the changes in population, industry, and air pollution levels that have occurred in Steubenville as well as the city's historical role as a $PM_{2.5}$ research center for the industrialized eastern United States, a reexamination of fine particle air pollution in the Steubenville region is warranted before enforcement of the NAAQS for $PM_{2.5}$.

The Steubenville Comprehensive Air Monitoring Program (SCAMP) was conducted from May 2000 to May 2002 to provide this examination. The study included two interdependent components. $PM_{2.5}$ and co-pollutants were sampled in the personal breathing space and homes of children and elderly volunteers. Additionally, $PM_{2.5}$ and co-pollutants were sampled outside the subjects' homes, at a central outdoor urban site in Steubenville, and at outdoor satellite sites located at the four cardinal compass points around this central site.

This paper is the first in a series of four papers that will present results from the central Steubenville site and

the four satellite sites. Its objectives are 2-fold. First, the paper provides an overview of the outdoor ambient air monitoring portion of SCAMP and characterizes concentrations of $PM_{2.5}$ and its major ionic and carbonaceous components measured during the program. This characterization of $PM_{2.5}$ in the Steubenville region provides the foundation upon which future papers in the series will build. Second, when analyzing spatial variability and interrelationships among $PM_{2.5}$ and its components, this paper explores the use of appropriate time series statistical techniques that account for the autocorrelated nature of $PM_{2.5}$ data. Many $PM_{2.5}$ statistical analyses have failed to consider autocorrelation; the implications of this oversight are assessed.

METHODS

Site Locations and Sampling Program

The five sites comprising the SCAMP outdoor ambient air monitoring network (Figure 1) were selected to investigate the effects of distance, direction, and urban versus rural location on $PM_{2.5}$ concentration and composition.

The central site in Steubenville (ST) served as an outdoor air monitoring "super site" for the program. Mass concentrations of $PM_{2.5}$ and PM with an aerodynamic diameter less than or equal to $10 \mu\text{m}$ (PM_{10}) were measured daily from May 13, 2000, through May 14, 2002, according to the Federal Reference Methods (FRMs) for $PM_{2.5}$ and PM_{10} . The sampled filters were analyzed every fourth day (beginning May 16, 2000) to determine the mass concentrations of ionic species, including ammonium (NH_4^+), sulfate (SO_4^{2-}), nitrate (NO_3^-), and chloride (Cl^-), as well as the water-extractable mass concentrations of 21 elements. A speciation sampler collected additional $PM_{2.5}$ samples every fourth day from August

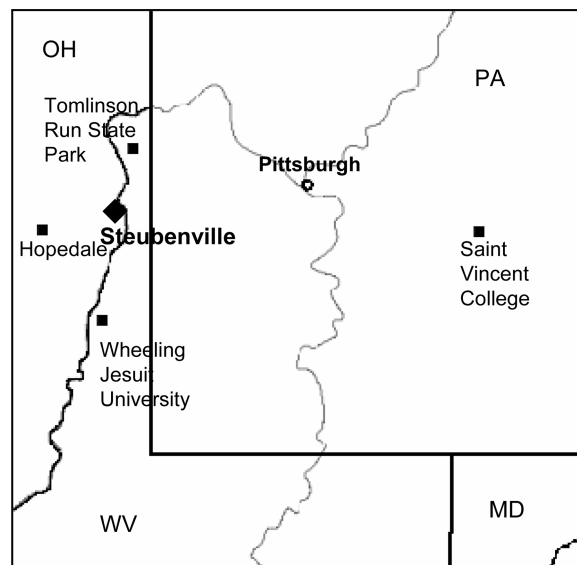


Figure 1. Map of SCAMP outdoor ambient air monitoring sites.

12, 2000, through May 14, 2002, which were analyzed to determine mass concentrations of elemental carbon (EC) and organic carbon (OC), as well as the acid-digestible concentrations of 21 elements. Additionally, $PM_{2.5}$, gaseous co-pollutants (SO_2 , NO_x , CO, O_3), and meteorological conditions were monitored continuously during the program, and pollen and mold spores were collected daily. The monitoring station was situated on the campus of Franciscan University of Steubenville atop a bluff overlooking the Ohio River at a height of ~ 110 m. The former Six Cities Study monitoring site had been located on the same campus ~ 0.3 miles southwest of the SCAMP site at a similar elevation. The campus is located in the northern part of Steubenville, and it is within a few miles of several major industrial facilities.

The four satellite sites were situated in comparatively rural areas. Relative to Steubenville, the sites were located 12 miles to the north at Tomlinson Run State Park (N) in New Manchester, WV; 14 miles to the west at Hopedale (W), OH; 20 miles to the south on the campus of Wheeling Jesuit University (S) in Wheeling, WV; and 67 miles to the east on the campus of Saint Vincent College (E) in Latrobe, PA. As illustrated in Figure 1, the Pittsburgh metropolitan area separates the eastern site from the other monitoring sites. The southern site was most similar to Steubenville, because it was located less than 2 miles from the center of Wheeling, a city of $\sim 30,000$ people located along the Ohio River. At each of the four satellite sites, $PM_{2.5}$ was collected daily from May 13, 2000, through May 14, 2002, according to the FRM, and its composition was analyzed every fourth day (beginning May 16, 2000) to determine the mass concentrations of ionic species and the water-extractable mass concentrations of 21 elements.

This paper focuses on the 24-hr average concentrations of $PM_{2.5}$, PM_{10} , ions, and carbon measured by the SCAMP outdoor ambient air monitoring network. Throughout SCAMP, the 24-hr sampling period for a given day occurred from 9:00 a.m. on that day until 9:00 a.m. on the following day to be consistent with the indoor and personal sampling program.

Experimental Methods

$PM_{2.5}$ and PM_{10} were collected and determined according to the FRMs for $PM_{2.5}$ and PM_{10} , respectively. Procedures were consistent with the guidelines outlined in the EPA Quality Assurance Guidance Document 2.12.¹⁵

Concentrations of NH_4^+ , NO_3^- , SO_4^{2-} , and Cl^- were determined by ion chromatography from samples collected on FRM filters. The water-soluble components of the PM sample were extracted ultrasonically from a Teflon filter with deionized water; anionic species and NH_4^+ were determined from separate aliquots of the leachate. Because the actual raw readings for determinations below

the detection limits of the instruments were not available, these determinations were set equal to half of the detection limit value. Detection limit values corresponded to ambient concentrations of $\sim 0.02 \mu\text{g}/\text{m}^3$ for Cl^- , $0.06 \mu\text{g}/\text{m}^3$ for NO_3^- , and $0.1 \mu\text{g}/\text{m}^3$ for NH_4^+ and SO_4^{2-} . Mass concentrations of Cl^- determined from exposed FRM filters are reported in this paper; however, because Cl^- contamination is difficult to control and Cl^- accounts for a very small portion of total $PM_{2.5}$ mass in the Steubenville region (many Cl^- determinations were below the detection limit), no trends or conclusions are drawn from these data.

At Steubenville, carbon species were collected every fourth day on a quartz filter with a Thermo Andersen $PM_{2.5}$ Speciation Sampler (RAAS2.5-400; Andersen Instruments Inc., Smyrna, GA), with the carbon channel operating at a flow rate of 7.5–8.5 L/min. No denuder was used in series with the quartz filter. Before use, the quartz filters were prepared by firing to 600°C for 4 hr to remove carbon background. Separate field blanks corresponding to each exposed filter were used to assess the carbon background experienced during handling and storage. To determine the 24-hr average concentrations of EC and OC, a square of known area was punched from the exposed filter and analyzed with a Sunset Laboratory Thermal Optical Transmittance (TOT) Analyzer (Sunset Laboratory Inc., Tigard, OR). The procedure used a thermal program similar to that followed by the EPA Speciation Trends Network,¹⁶ but it was modified slightly in accordance with the instrument manufacturer's recommendations. For each exposed filter, results were adjusted to account for background carbon by subtraction of the corresponding field blank. Daily average concentrations of EC and OC were then calculated based on the masses determined by this method and the volume of air sampled. Additionally, the concentration of OC was multiplied by a factor of 1.4 to estimate the concentration of organic material (OM). Turpin and Lim¹⁷ suggest a factor of 1.4–1.8 for approximating the average molecular weight per carbon weight for urban organic aerosols. Lacking additional information specific to Steubenville, we used the conventional low-end value of 1.4 for this analysis to avoid overstating the mass contribution of OM.

It is important to recognize that the FRMs for $PM_{2.5}$ and PM_{10} do not employ denuders to scrub reactive gases, back-up filters to collect revolatilized material, or blank correction methods to adjust for exposure during handling and storage. Thus, concentrations of PM and its components measured by these methods can deviate from the true ambient concentrations. Biases for semi-volatile species, including NO_3^- , NH_4^+ , and OC, have been documented.¹⁸ As an example of bias in the FRM,

NO_3^- values determined from FRM Teflon filters at Steubenville were on average $0.79 \mu\text{g}/\text{m}^3$ lower than NO_3^- values determined from samples collected on nylon filters via a collocated speciation sampler equipped with a magnesium oxide denuder. This underestimation represents 39% of the overall average NO_3^- mass concentration determined by the latter method. Nevertheless, because compliance with the NAAQS for $\text{PM}_{2.5}$ is based upon FRM-determined mass concentrations, the mass concentrations of all species other than EC and OM reported in this paper were determined from samples collected according to the FRM. Analytical limitations prohibit the determination of EC and OM from FRM Teflon filters; however, these species were collected without a denuder to better approximate FRM sampling conditions.

Statistical Methods

A number of studies¹⁹⁻²¹ assessing the spatial variability of PM or interrelationships among particulate components have used Pearson correlation coefficients (r) or simple linear regression analysis coefficients of determination (R^2 ; $R^2 = r^2$) to describe and compare strengths of association between variables. Simple t tests have often been employed to determine statistical significance. Pinto et al.²² recently criticized studies that relied solely on the use of Pearson correlations to assess spatial variability, arguing that these correlations should be reported alongside a statistic that measures uniformity in the magnitude of concentrations. These studies, however, have generally failed to address important statistical issues, most notably autocorrelation, that are relevant to air monitoring time series data.

Pearson correlations, simple linear regression analyses, and t tests only provide meaningful results if the assumptions upon which they are based are reasonably satisfied. In addition to requiring a correctly specified linear model, linear regression assumes that the error component is normally distributed, statistically independent (random), and homoskedastic (has a constant variance). A paired t test assumes that the differences between the two variables under consideration are randomly sampled from a normally distributed population. Raw ambient air pollutant concentration data frequently fail to meet these assumptions.

It is most serious that the assumption of randomness (statistical independence) is violated because of serial or autocorrelation. For a time series of observations, the lag- k autocorrelation coefficient is the correlation between the observations in the series and the corresponding observations in the same series occurring k time intervals earlier. Granger and Newbold²³ discuss the possible occurrence of spurious correlations when serial dependencies are not

appropriately handled, and Milionis and Davies²⁴ specifically emphasize the need to account for autocorrelation when analyzing air pollution data. Time series of daily data may exhibit autocorrelation at 1- or several-day lags, as well as at weekly and seasonal lags. If not properly addressed, serial correlation can lead to an incorrect conclusion about the existence or strength of a relationship between two time series. As a simple example, NO_3^- concentrations in the Steubenville region exhibit a pronounced seasonal pattern. If serial correlation is not considered, it is difficult to determine whether a correlation observed between NO_3^- concentrations measured at two distinct sites is truly indicative of an association between day-to-day fluctuations in NO_3^- levels at the sites or rather is simply induced by the predominance of the seasonal pattern.

Thus, before performing linear regression analysis to study $\text{PM}_{2.5}$ and component data collected during SCAMP, the time series were transformed by the natural logarithm, square root, or fourth root, to render their error distributions more normal. Such transformations also often help to improve homoskedasticity. For each transformed variable, serial dependencies were identified in the standard way by examining the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots. Where necessary, autocorrelation at early lags was modeled with an autoregressive integrated moving average (ARIMA) model.²⁵ The general form of an ARIMA (p, d, q) model representing a time series, X , of appropriately transformed air pollutant concentrations is

$$W_t = \eta + \phi_1 W_{t-1} + \phi_2 W_{t-2} + \dots + \phi_p W_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q} \quad (1)$$

where

$$W_t = \nabla^d X_t \quad (2)$$

Here, ∇ is the difference operator and d is the order of differencing such that for $d = 0$, $W_t = X_t$, and for $d = 1$, $W_t = X_t - X_{t-1}$. X_t and ϵ_t are the transformed air pollutant concentration and random error term, respectively, at time t ; $\phi_1, \phi_2, \dots, \phi_p$ are the autoregressive parameters; $\theta_1, \theta_2, \dots, \theta_q$ are the moving average parameters; and η is the intercept term.

When estimating the ARIMA model parameters, missing values were handled via a state-space representation of the process.²⁶ Seasonal patterns and other nonstationarities were removed by the inclusion of first-order differencing in the ARIMA model or by subtracting out a 3-month moving average before modeling, when appropriate. A suitable model was selected by minimizing the

Akaike information criterion (AIC), provided that the estimated parameters were reasonable and statistically significant. In all analyses presented in this paper, statistical significance was determined at a significance level (α) of 0.05. Residuals from the model were examined with an ACF plot and Ljung-Box tests for randomness to verify their statistical independence. The residuals obtained by transforming and modeling (where necessary) a series according to this procedure are henceforth referred to as the "adjusted" values of that series. These adjusted values were used in the simple linear regression analyses.

Examples of the transformation and form of the time series model applied to $PM_{2.5}$ and major component variables measured during SCAMP are shown in Table 1. The effect of adjusting data according to the ARIMA modeling procedure described above is demonstrated in Figure 2, which shows ACF plots for log-transformed NO_3^- data measured at the northern satellite site before (a) and after (b) adjustment with an ARIMA model. Figure 2a reveals that even though NO_3^- was determined on a less-than-daily (1-in-4-day) frequency, the NO_3^- time series exhibited statistically significant autocorrelation at early lags. ARIMA modeling was successful in removing this statistically significant autocorrelation, as shown in Figure 2b. The removal of autocorrelation caused a 53% reduction in the variance of the data, indicating that much of the variability in the original time series could be accounted for by serial correlation.

Figure 3 evaluates the randomness of residuals from simple linear regression of eastern site NO_3^- versus northern site NO_3^- with log-transformed (a and c) and adjusted (b and d) data sets. Residuals from regression analysis with log-transformed data that were not ARIMA-adjusted exhibited small but statistically significant autocorrelations at 4- and 8-day lags (a); Ljung-Box test results (c) confirmed that these residuals are not random at $\alpha = 0.05$. The use of ARIMA-adjusted data sets eliminated the problem of autocorrelated residuals (b) and satisfied the assumption of randomness (d). When ARIMA-adjusted rather than log-transformed data were used in the simple linear regression analysis, R^2 decreased from 0.57 to 0.30, and the P -value testing the hypothesis that the slope of the regression line is equal to zero increased by 4 orders of magnitude, although it remained less than 0.0001.

Despite these efforts, the residuals of several regression analyses with ARIMA-adjusted data still exhibited a minor amount (on the order of $r = 0.2$) of statistically significant autocorrelation at an early lag. For the purposes of this paper, given practical limitations, no further adjustments were made in these instances. However, to better handle these cases, a single model with moving average parameters fitted to the autocorrelated error structure and simultaneous estimation of all regressive and autoregressive parameters

Table 1. Transformation and form of time series model applied to $PM_{2.5}$ and ionic component variables.

Location/Species	Transformation	Moving Average ^b	ARIMA Parameters ^a		
			<i>p</i>	<i>d</i>	<i>q</i>
Steubenville					
PM _{2.5}	ln(x)	—	1	1	1,2
NH ₄ ⁺	ln(x)	—	—	—	—
NO ₃ [−]	ln(x)	—	0	1	1
SO ₄ ^{2−}	ln(x)	21	—	—	—
Other ^c	ln(x)	21	—	—	—
North					
PM _{2.5}	ln(x)	—	0	1	1–3
NH ₄ ⁺	ln(x)	—	1	0	0
NO ₃ [−]	ln(x)	—	0	1	1
SO ₄ ^{2−}	ln(x)	21	—	—	—
Other ^c	ln(x)	—	—	—	—
South					
PM _{2.5}	ln(x)	—	1	1	1
NH ₄ ⁺	ln(x)	—	0	0	1
NO ₃ [−]	ln(x)	—	3	1	1
SO ₄ ^{2−}	ln(x)	21	0	0	2
Other ^c	ln(x)	—	—	—	—
East					
PM _{2.5}	ln(x)	—	0	1	1–3
NH ₄ ⁺	ln(x)	—	3	0	0
NO ₃ [−]	ln(x)	—	1	0	1
SO ₄ ^{2−}	ln(x)	21	0	0	1
Other ^c	ln(x)	—	0	0	1
West					
PM _{2.5}	ln(x)	—	6,7,18	1	1,2
NH ₄ ⁺	ln(x)	—	—	—	—
NO ₃ [−]	ln(x)	—	0	1	1
SO ₄ ^{2−}	ln(x)	21	—	—	—
Other ^c	ln(x)	—	0	0	1

^aFor an ARIMA (p, d, q) model, p and q indicate the lags (in units of 1 day for $PM_{2.5}$, 4 days for components) to which autoregressive and moving average parameters were fit, and d indicates the order of differencing used in the ARIMA model; ^bIndicates that a moving average of the specified number of measurements was subtracted from the transformed data to stationarize the series; ^cConcentrations of 'other' components are by difference and include concentrations of EC and OM.

would be required. Such a model was applied to an example case in which $PM_{2.5}$ values measured at the northern site were regressed against those measured at Steubenville. Use of the more sophisticated model did not substantially alter the results. For simple linear regression with adjusted values, the slope was 0.69 ± 0.02 ; the slope estimated by the more sophisticated model was 0.68 ± 0.02 . For both models, the intercept was not significantly different from zero.

When examining spatial uniformity in the magnitude of $PM_{2.5}$ and component concentrations, paired t

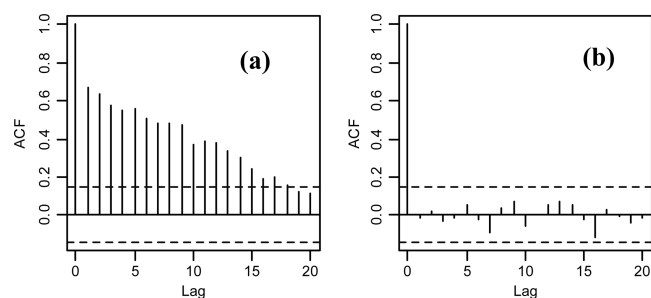


Figure 2. Autocorrelation function plots for log-transformed NO_3^- concentration data measured at the northern (N) satellite site (a) before and (b) after adjustment with an ARIMA (0, 1, 1) model. Lags are in units of 4 days. Lines extending beyond the dashed interval indicate statistically significant correlations.

tests were required to test whether concentrations measured at Steubenville were significantly greater than those measured at satellite sites. Differences between Steubenville and satellite site concentration pairs were not normally distributed; hence, differences between log-transformed concentrations were used. The tests, then, ultimately examined whether the mean ratios of the Steubenville concentrations to the satellite site concentrations differed significantly from one. ARIMA models were fit to those series of differences that exhibited serial correlation. The intercept term of the ARIMA model is a function of the mean difference μ such that $\mu = \eta / (1 - \phi_1 - \phi_2 - \dots - \phi_p)$. Values of μ and its standard error (SE) were estimated as part of the ARIMA modeling process; SE

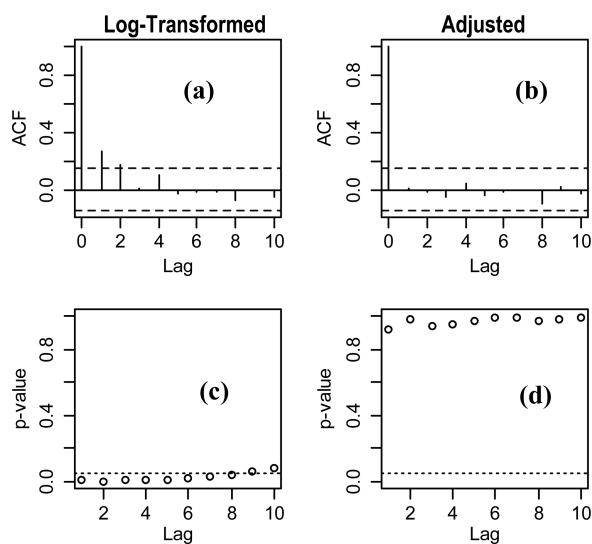


Figure 3. Analysis of residuals from simple linear regression of eastern (E) site NO_3^- concentrations as a function of northern (N) site NO_3^- concentrations with (a) and (c) log-transformed data and (b) and (d) adjusted data. Lags are in units of 4 days. Plots (a) and (b) are autocorrelation function plots of the residuals; lines extending beyond the dashed intervals indicate statistically significant autocorrelations. Plots (c) and (d) are p -values for the Ljung-Box statistic for various lags of the residuals; the dotted lines in these plots denote $p = 0.05$.

computed in this way provides an estimate of the uncertainty that accounts for the presence of autocorrelation. Thus, the t tests were performed with these ARIMA-estimated values. p -Values resulting from tests performed with this time series procedure were often substantially greater than p -values resulting from simple paired t tests applied to raw data, indicating that the use of these simple t tests yields misleading conclusions about statistical significance for autocorrelated observations. For example, a simple paired t test performed on the differences between SO_4^{2-} concentrations measured at the Steubenville and western sites yielded a p -value of 3.1×10^{-7} ; when the modified procedure was used, the p -value increased to 2×10^{-3} .

RESULTS AND DISCUSSION

Overall PM and Component Concentrations

Descriptive statistics for all valid 24-hr average concentrations of $\text{PM}_{2.5}$ and its ionic and carbonaceous components measured at the five monitoring sites between May 13, 2000, and May 14, 2002, are presented in Table 2. $\text{PM}_{2.5}$ concentrations ranged from less than 3 to greater than $50 \mu\text{g}/\text{m}^3$ at each of the five sites, with coefficients of variation ranging from 54.9% to 61.1%.

The mean 24-hr average $\text{PM}_{2.5}$ concentration measured at the Steubenville SCAMP site during the year 2001 ($18.1 \mu\text{g}/\text{m}^3$) differed by only $0.1 \mu\text{g}/\text{m}^3$ from the 2001 mean concentration ($18.2 \mu\text{g}/\text{m}^3$) reported by the Ohio EPA²⁷ for a monitor located in central Steubenville ~1 mile south of the SCAMP site, and by $0.8 \mu\text{g}/\text{m}^3$ from the 2001 mean concentration ($18.9 \mu\text{g}/\text{m}^3$) reported for a monitor located in Mingo Junction ~4 miles south of the SCAMP site. This suggests that $\text{PM}_{2.5}$ concentrations measured atop the bluff at the SCAMP site are, on average, representative of the magnitude of concentrations experienced throughout the Steubenville urban area. The overall mean $\text{PM}_{2.5}$ concentration measured at Steubenville during SCAMP was $18.4 \mu\text{g}/\text{m}^3$. This result indicates that average ambient $\text{PM}_{2.5}$ concentrations in Steubenville have decreased by $11.2 \mu\text{g}/\text{m}^3$, or by 37.8%, since Six Cities Study investigators associated Steubenville's $\text{PM}_{2.5}$ concentrations with mortality by using $\text{PM}_{2.5}$ data collected between 1979 and 1985.¹ The substantial long-term decrease in $\text{PM}_{2.5}$ concentrations is likely attributable to advances in the application of air pollution control technology and to the decline of steel and other manufacturing industries in Steubenville. Nevertheless, the mean $\text{PM}_{2.5}$ concentration measured at Steubenville during SCAMP exceeded the annual NAAQS for $\text{PM}_{2.5}$ by $3.4 \mu\text{g}/\text{m}^3$. In fact, more than half of all valid $\text{PM}_{2.5}$ concentrations measured at Steubenville were greater than $15 \mu\text{g}/\text{m}^3$. Although based on only 2 yr rather than the required 3 yr of data, these findings suggest that a

Table 2. Summary statistics for PM and component mass concentrations (in $\mu\text{g}/\text{m}^3$) measured at SCAMP sites, 2000–2002.

FRM Mass Concentration														
Site	PM Size Fraction	N	Mean	CV (%) ^a	Median	Range	P98 ^b	SO ₄ ²⁻	NO ₃ ⁻	NH ₄ ⁺	Cl ⁻	Other ^c	EC	OM
ST	2.5	640	18.4	61.1	15.3	2.7–64.8	49.1	5.8	1.16	2.3	0.21	9.0	0.84 ^d	4.5 ^d
		151						(23.4)	(5.34)	(7.2)	(3.76)	(26.2)	(4.62)	(15.3)
N	2.5	663	13.9	55.1	12.1	2.5–54.5	34.4	4.5	0.71	1.7	0.06	6.3	NA	NA
		161						(33.2)	(4.80)	(7.6)	(0.79)	(20.4)		
S	2.5	416	16.0	54.9	13.6	2.8–51.1	39.9	5.9	0.73	2.2	0.07	7.4	NA	NA
		96						(21.8)	(4.59)	(7.3)	(0.36)	(28.4)		
E	2.5	672	15.2	57.2	13.0	2.8–55.1	40.6	5.1	0.74	1.9 ^e	0.07	7.1	NA	NA
		155						(28.4)	(4.42)	(6.9)	(2.04)	(28.8)		
W	2.5	587	13.9	59.7	11.9	2.3–64.8	36.5	4.9	0.88	1.8	0.05	6.1	NA	NA
		129						(31.6)	(4.93)	(6.1)	(0.27)	(23.3)		
ST	10	676	27.4	58.0	23.1	4.9–102.2	67.5	6.2	1.76	2.2	0.38	17.1	NA	NA
		161						(31.8)	(6.55)	(8.5)	(4.67)	(53.8)		
ST	10–2.5 ^f	610	9.0	82.6	7.5	–19.6–47.9	32.8	0.3	0.61	–0.1	0.20	8.9 ^g	NA	NA
		141						(5.3)	(2.10)	(1.3)	(3.11)	(30.1)		

Note: For N, the top value is the number of valid FRM mass concentration data; the bottom value is the number of valid PM component data (except where noted otherwise); for PM components, the top value is the mean concentration; the bottom value (in parentheses) is the maximum concentration. ^aCV = coefficient of variation = $100 \times \text{mean}/\text{standard deviation}$; ^bP98 = 98th percentile; ^cConcentrations of 'other' were computed by difference and include the concentrations of elemental carbon and organic material; ^dN = 142; ^eN = 158; ^fConcentrations of PM_{10–2.5} and its ionic components were computed by differencing PM₁₀ and PM_{2.5} data; ^gN = 134.

PM_{2.5} reduction strategy will likely be required for Steubenville. Mean and median concentrations of PM_{2.5} measured at each of the four satellite sites were lower than those measured at Steubenville; however, the southern and eastern sites still exceeded the annual PM_{2.5} standard, with mean concentrations of 16 and 15.2 $\mu\text{g}/\text{m}^3$, respectively. The northern and western sites each had a mean concentration of 13.9 $\mu\text{g}/\text{m}^3$.

No PM_{2.5} concentration measured at any of the five monitoring sites during SCAMP exceeded the 24-hr standard of 65 $\mu\text{g}/\text{m}^3$, although the Steubenville and western sites each registered a maximum 24-hr average concentration of 64.8 $\mu\text{g}/\text{m}^3$. Ninety-eighth percentile concentrations at the five sites were well in compliance, ranging from 15.9 to 30.6 $\mu\text{g}/\text{m}^3$ below the 24-hr standard.

Sulfate, nitrate, and ammonium all contribute substantially to the mass of PM_{2.5} in the Steubenville region. Sulfate is the most prevalent of these ionic components, with mean 24-hr concentrations of 4.5–5.9 $\mu\text{g}/\text{m}^3$ and maximum concentrations of 21.8–33.2 $\mu\text{g}/\text{m}^3$ measured at the five SCAMP monitoring sites. Table 3 summarizes the ionic composition of PM_{2.5} measured at each of the sites. On average, the PM_{2.5} composition was fairly homogeneous regionally, with the four ionic species accounting for more than half the total mass. However, the composition experienced on a particular day could deviate appreciably from this average, as evidenced by the ranges listed in the table. The almost 50% of PM_{2.5} mass not accounted for by the four ions is likely composed of carbon species, crustal material, trace elements, and some

Table 3. Ionic composition of PM_{2.5} measured at the SCAMP monitoring sites.

Location	wt% Ammonium	wt% Nitrate	wt% Sulfate	wt% Chloride	wt% Other ^a
Steubenville	12.5 (8.3, 16.4)	6.3 (1.3, 18.0)	31.3 (19.7, 41.0)	1.1 (0.06, 2.2)	48.8 (34.6, 62.4)
Northern site	13.0 (8.7, 15.9)	5.4 (0.5, 15.5)	34.1 (19.7, 42.9)	0.4 (0.06, 1.3)	47.1 (34.3, 63.7)
Southern site	13.4 (8.8, 16.3)	4.5 (0.4, 16.6)	36.2 (19.0, 46.8)	0.4 (0.05, 1.3)	45.5 (34.4, 63.5)
Eastern site	12.9 (7.8, 16.5)	4.9 (0.6, 14.9)	34.2 (18.3, 44.3)	0.4 (0.05, 1.3)	47.6 (35.7, 66.8)
Western site	13.0 (8.6, 16.3)	6.4 (0.4, 18.6)	35.7 (20.9, 47.3)	0.3 (0.06, 1.1)	44.6 (31.6, 64.6)

Note: For each component/location pair, the first number is the overall average weight percent of the component at that location. The numbers in parentheses are the 10th and 90th percentile daily weight percents. Where concentrations were below detection limit values, weight percentages were calculated assuming half of the detection limit concentration. ^aComputed by difference; includes EC and OM.

water. Measurements of EC and OC at Steubenville confirmed their significance as components of $PM_{2.5}$. On average, EC accounted for 4.7% and OM accounted for 25% of the total $PM_{2.5}$ mass, leaving ~19% unaccounted for. Again, the weight percentages of these carbonaceous components were variable from day to day, ranging from a 10th percentile of 1.9% to a 90th percentile of 8.3% for EC, and from a 10th percentile of 13.5% to a 90th percentile of 40.1% for OM. The high average relative abundance of fine OM, which consists of hundreds to thousands of different components that have not been well characterized,²⁸ emphasizes the need for additional research on the health effects of the components of $PM_{2.5}$.

Table 2 also summarizes the concentrations of PM_{10} and its components measured at Steubenville, as well as the contribution of coarse PM ($PM_{10-2.5}$) and its components to these concentrations. Concentrations of $PM_{10-2.5}$ and its components were computed by difference. Negative values reported for the coarse fraction indicate the presence of measurement errors in the FRMs. On average (geometric mean), $PM_{2.5}$ accounted for ~66% of the mass of PM_{10} at Steubenville. Nitrate and especially sulfate and ammonium tended to be more prevalent in the fine fraction than in the coarse fraction. The relatively large unexplained portion of the coarse fraction likely includes an abundance of dust and crustal materials emitted by primary sources, as well as some carbonaceous species and trace elements.

Temporal Variability

Temporal patterns among $PM_{2.5}$ and its components are now studied to gain further insight into the nature of $PM_{2.5}$ in the Steubenville region. Despite exhibiting statistically significant autocorrelation at 1- and several-day lags, $PM_{2.5}$ concentrations in the Steubenville region exhibited appreciable day-to-day variability. The average magnitude of the day-to-day change in $PM_{2.5}$ concentration ranged from 5.3 to 7.6 $\mu\text{g}/\text{m}^3$ at the five monitoring sites, with daily positive and negative changes of 40 $\mu\text{g}/\text{m}^3$ or more observed in several instances.

Seasonal patterns in $PM_{2.5}$ concentrations are examined in Figure 4. The box plots in Figure 4a show that $PM_{2.5}$ concentrations of 40 $\mu\text{g}/\text{m}^3$ or more could occur at Steubenville during any season of the year. Based on the monthly averages plotted in Figure 4b, $PM_{2.5}$ at all five sites exhibited a pronounced seasonal trend of higher summertime and lower wintertime concentrations. Mean August concentrations at the five sites ranged from 19.2 to 27 $\mu\text{g}/\text{m}^3$, whereas mean December concentrations ranged from 9.3 $\mu\text{g}/\text{m}^3$ to 13.1 $\mu\text{g}/\text{m}^3$.

To better understand the nature of the observed seasonal variations in $PM_{2.5}$, seasonal patterns in the concentrations of its major components were studied. Figure 5

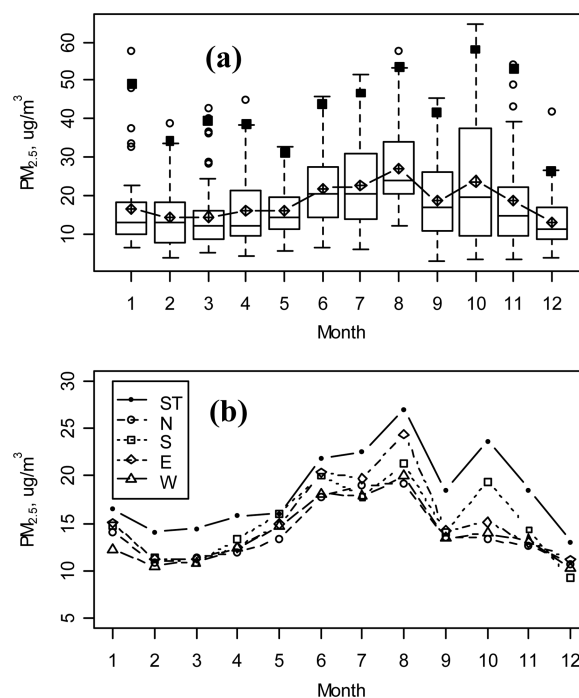


Figure 4. Seasonal patterns among $PM_{2.5}$ concentrations measured between May 13, 2000, and May 14, 2002. Plot (a) shows the distribution of $PM_{2.5}$ concentrations by month at Steubenville. The horizontal lines denote median concentrations; the boxes extend to the 25th and 75th quartiles; the whiskers extend to the most extreme data points within 1.5 times this interquartile range. Circular points represent outliers, solid squares denote 98th percentile concentrations, and connected diamonds represent mean concentrations. Plot (b) shows mean $PM_{2.5}$ concentrations at all five monitoring sites.

shows mean monthly concentrations of major ionic and carbonaceous components measured at Steubenville. Although not shown, seasonal patterns in the concentrations of SO_4^{2-} , NO_3^- , and NH_4^+ measured at the four

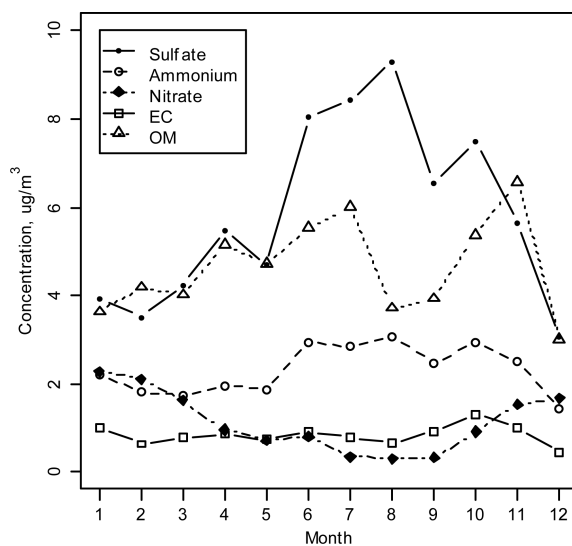


Figure 5. Mean monthly concentrations of major ionic and carbonaceous components of $PM_{2.5}$ measured at Steubenville during SCAMP.

satellite sites were generally similar to those observed at Steubenville. Noteworthy seasonal trends are evident among the SO_4^{2-} and NO_3^- concentrations, with the highest SO_4^{2-} concentrations occurring during the summer and into the fall, and the highest NO_3^- concentrations occurring during the winter. Ammonium, which is generally assumed to be present as $(\text{NH}_4)_2\text{SO}_4$, NH_4HSO_4 , or NH_4NO_3 , exhibits a seasonal pattern somewhat similar to that of SO_4^{2-} . These seasonal patterns are typical of the northeastern United States²⁹ and are generally explainable by seasonal variations in meteorological conditions. Increased summertime photochemical activity likely accounts at least in part for the observed elevated summertime SO_4^{2-} concentrations.³⁰ The formation of particulate nitrate is thermodynamically favored by low temperatures and high relative humidity and depends on the availability of ammonia (NH_3),³¹ explaining why NO_3^- concentrations are elevated during the winter when less NH_3 has been scavenged to form ammonium sulfate. Whereas the observed correspondence between elevated $\text{PM}_{2.5}$ concentrations and elevated SO_4^{2-} concentrations during the summer months may seem to suggest that reductions in sulfate will achieve corresponding reductions in $\text{PM}_{2.5}$, research in the southeastern United States³² has suggested that resulting available NH_3 will instead react to form particulate nitrate, thereby lessening the effectiveness of this strategy. Although their average concentrations varied from month to month, EC and OM did not exhibit discernible seasonal cycles.

Concentrations of $\text{PM}_{2.5}$ and of its ionic components did not display substantial day-of-the-week trends. Trends in EC and OM at Steubenville were more pronounced, as illustrated in Figure 6. Concentrations generally built through the workweek from Sunday through Thursday and then declined to near-initial levels on Friday and Saturday. This suggests the influence of some currently unknown anthropogenic emission source that follows a similar weekly pattern.

Spatial Variability

Linear regression analysis was used to study whether daily $\text{PM}_{2.5}$ concentrations measured at the five SCAMP monitoring sites were correlated. As described previously, the analyses were performed with adjusted data sets, so that the reported relationships are not influenced by the effects of serial correlation present in each raw data set. In all cases, the removal of autocorrelation via ARIMA modeling caused a reduction in the apparent strength of the relationship between $\text{PM}_{2.5}$ concentrations at any two sites being compared. However, even after accounting for early-lag autocorrelation and seasonal trends, daily adjusted $\text{PM}_{2.5}$ values for all possible site pairings exhibited

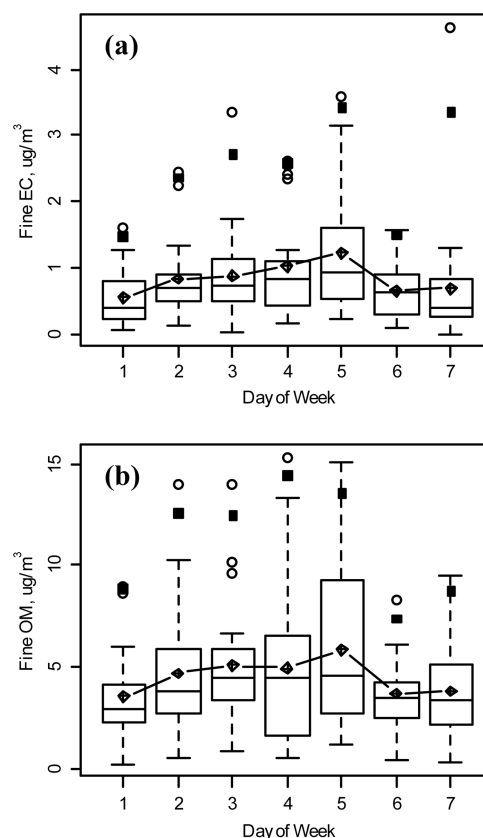


Figure 6. Distribution of (a) EC and (b) OM concentrations at Steubenville by day of the week, based on measurements taken every fourth day between August 12, 2000, and May 14, 2002. Box plots are constructed as in Figure 4a. Sunday is day 1.

positive, statistically significant ($p < 0.001$) linear relationships. Figure 7 plots R^2 values describing these site-to-site relationships among adjusted $\text{PM}_{2.5}$ data as a function of distance between the sites. These results indicate

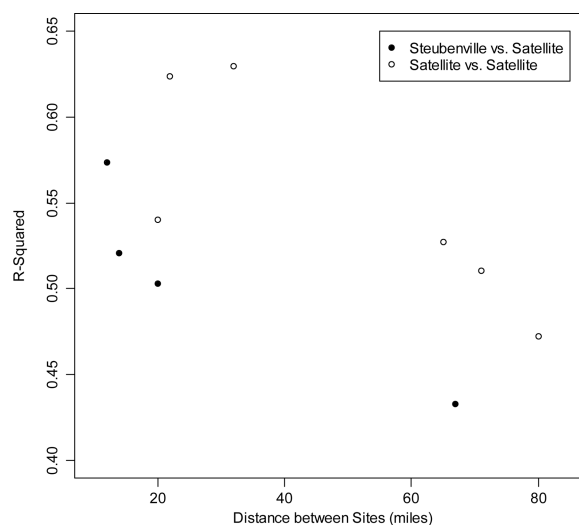


Figure 7. R^2 values from simple linear regression analyses performed on $\text{PM}_{2.5}$ data for all possible SCAMP site pairs. R^2 values are plotted against intersite distances. Data were adjusted before regressing to normalize them and remove serial correlation.

that between 43% and 63% of the variance in the adjusted $PM_{2.5}$ values at one site can be accounted for by the adjusted $PM_{2.5}$ values at any of the other four monitoring sites. The plot also reveals a general decrease in the strength of the relationship between sites as the distance separating them increases, although this trend is only manifest when the full domain of distances is considered. It is possible that the trend is due to the influence of emissions from the Pittsburgh metropolitan area on the most remote eastern site rather than to distance itself, because it is only observed when relationships between the eastern site and the sites west of Pittsburgh are considered. Finally, at comparable distances of separation, it appears that satellite sites tend to be better correlated with each other than with Steubenville.

Similar regression analyses were performed for all site pairs with adjusted sulfate, nitrate, ammonium, and "other" (by difference) data. Again, even after removing serial correlation from each data set, adjusted values for each of these components exhibited positive, statistically significant ($p < 0.01$) linear relationships for all possible site pairings. Figure 8 shows R^2 values obtained from the regression analyses for each component plotted as a function of the distance separating the sites being compared. Sulfate, ammonium, and the "other" component displayed rather strong spatial relationships, with the regression models accounting for 33% to 88% of the observed

variance. Site-to-site NO_3^- relationships were not as pronounced, with R^2 values ranging from 0.11 to 0.54. For all of the components, a decrease in the strength of the intersite relationship was observed with increasing distance. Again, the influence of the Pittsburgh metropolitan area may account at least in part for the observed trend.

These results indicate that $PM_{2.5}$ and several of its major components display an appreciable amount of regional homogeneity. Concentrations at distinct locations tend to vary together to some extent over and above variations that are induced by serial correlation. In general, for sites separated by up to 30 miles, distance has no obvious effect on the strength of the relationship between the two sites. As the separation distance is increased to 60–80 miles, a downward trend in strength is observed. However, it is important to note that the eastern site, despite being separated from the other four sites by over 60 miles and by the Pittsburgh metropolitan area, nonetheless exhibits statistically significant positive correlations with each of these sites for $PM_{2.5}$ and for the four components that were studied. These results are consistent with previous studies^{33,34} that provided evidence of the influence of long-range transport and regional meteorology on variations in $PM_{2.5}$ concentrations in the eastern United States. The effect of meteorology on $PM_{2.5}$ concentrations measured during SCAMP is examined in the second and third papers in this series.

The results obtained here according to statistical procedures that account for the autocorrelated nature of $PM_{2.5}$ data are also generally consistent with the findings of previous studies that did not account for autocorrelation. On the basis of simple graphical comparisons, Robinson et al.³⁵ reported that in the Pittsburgh area, $PM_{2.5}$ concentrations measured at five distinct monitoring sites were well correlated. Sulfate concentrations were more strongly associated across the sites than total $PM_{2.5}$ mass concentrations, whereas variations in nitrate concentrations were less regionally homogeneous. Hansen et al.²¹ reported similar findings for the southeastern United States on the basis of simple Pearson correlations. Pinto et al.²² reported that $PM_{2.5}$ mass concentrations at four monitoring sites in the Steubenville-Weirton MSA were well correlated for all site pairs ($r = 0.8$ – 0.88), whereas correlations among site pairs in the Pittsburgh MSA, although appreciable, spanned a wider range ($r = 0.58$ – 0.94). The agreement between these previous findings and the findings presented in this paper suggest that simple descriptive statistics and correlation analyses applied to raw $PM_{2.5}$ data may be sufficient for drawing general conclusions about trends in the spatial variability of $PM_{2.5}$ and its major components. However, adjustment for autocorrelation was not a futile undertaking because it allowed autocorrelation to be ruled out as a potential

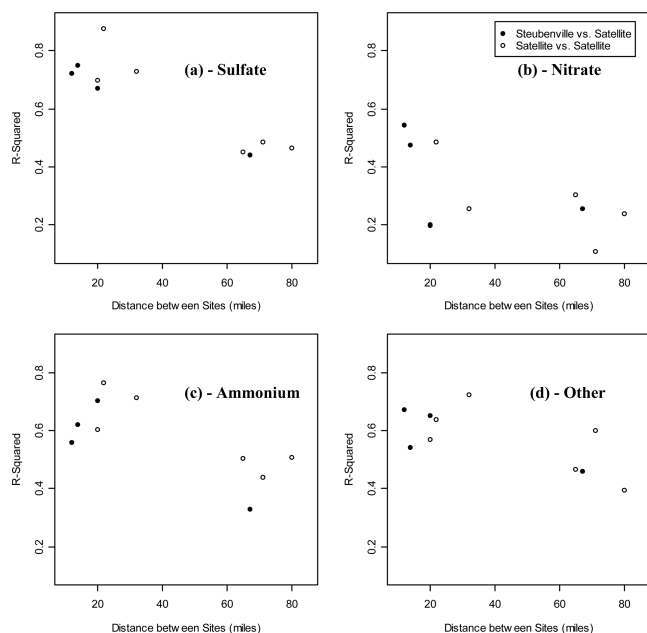


Figure 8. R^2 values from simple linear regression analyses performed on fine fraction (a) sulfate, (b) nitrate, (c) ammonium, and (d) "other" (by difference, includes EC and OM) data for all possible SCAMP site pairs. R^2 values are plotted against intersite distances. Data were adjusted before regressing to normalize them and remove serial correlation.

cause for the observed associations. Any study that has not considered autocorrelation risks the possibility that once autocorrelation is accounted for, the significance of the result will change drastically. Examples provided in the Statistical Methods section highlighted this possibility. Analyses attempting to assess statistical significance or to precisely quantify relationships among sites for comparative or predictive purposes must account for autocorrelation to obtain meaningful results.

The effect of urban versus rural location on $PM_{2.5}$ and component concentrations was studied via the modified paired t test procedure that has already been described. Table 4 presents the p -values obtained from these analyses. The mean raw differences between the Steubenville and satellite site concentrations and the percentage of days for which the Steubenville concentrations were greater than the satellite site concentrations are also provided. In all cases but one, $PM_{2.5}$ and component concentrations at Steubenville were significantly higher than concentrations at all four satellite sites, indicating the influence of local sources at Steubenville. The SO_4^{2-} concentration at the southern site, whose setting was most similar to Steubenville, did not differ significantly from that at Steubenville.

$PM_{2.5}$ and component concentrations measured at the northern and western sites are most representative of regional "background" concentrations in the Steubenville area. These sites were situated in relatively rural areas that are not generally directly downwind of emissions from Steubenville or Pittsburgh. The sites were located closer to Steubenville than the southern and eastern sites, but they were the only two SCAMP sites whose overall mean $PM_{2.5}$ concentrations did not exceed the annual NAAQS. Adjusted sulfate values for the northern and western sites were more highly correlated ($R^2 = 0.88$) than for any other site pair, suggesting that these sites are representative of regional background concentrations of secondary $PM_{2.5}$.

Assuming that the northern and western sites are regional background sites, the data in Table 4 suggest that local sources in Steubenville on average contribute $\sim 4.6 \mu\text{g}/\text{m}^3$ to the city's $PM_{2.5}$ concentration. Application of the same assumption to estimate the average local source contribution to $PM_{2.5}$ concentrations at the southern site yields an estimated contribution of $1.2 \mu\text{g}/\text{m}^3$, suggesting that local sources more heavily impacted Steubenville than the Wheeling site during SCAMP. The "other" component, whose major constituent ($\sim 60\%$ w/w) in Steubenville is carbonaceous material, accounts for more than half ($\sim 2.5 \mu\text{g}/\text{m}^3$) of the local contribution in Steubenville. This result is not surprising, as EC³⁶ and primary OM are expected to exhibit strong spatial concentration gradients, with higher concentrations observed near emission sources. Although SO_4^{2-} is predominantly a regional pollutant with average background concentrations of $\sim 4.5\text{--}5 \mu\text{g}/\text{m}^3$ in the Steubenville region, SO_4^{2-} concentrations measured at Steubenville were on average $\sim 1 \mu\text{g}/\text{m}^3$ higher than those measured at the background sites. However, the results collectively suggest that the reduction of $3.4 \mu\text{g}/\text{m}^3$ in average $PM_{2.5}$ concentration required for Steubenville to achieve compliance with the annual NAAQS cannot be achieved by reducing local emissions of any single $PM_{2.5}$ component in Steubenville; rather, a regional or multi-component strategy will likely be required.

Relationships among Components

To study interrelationships among components of $PM_{2.5}$, linear regression analyses were performed with adjusted data for pairs of components measured at Steubenville. Table 5 presents the results of these analyses. All of the intercomponent relationships are positive and statistically significant; however, those involving NO_3^- appear to be the weakest, even after accounting for the strong seasonal patterns that affect this variable. Statistically significant relationships among the sulfate, carbon, and "other" components, which have R^2 values between 0.2

Table 4. Comparison of $PM_{2.5}$ and component concentrations at Steubenville with those at the satellite sites.

	North	South	East	West
$PM_{2.5}$	<0.0001 (4.7, 83%)	<0.0001 (3.2, 70%)	<0.0001 (3.1, 64%)	<0.0001 (4.6, 83%)
SO_4^{2-}	<0.0001 (1.2, 72%)	0.27 (0.3, 59%)	0.015 (0.7, 56%)	0.0020 (1.1, 71%)
NH_4^+	0.0091 (0.5, 73%)	0.0031 (0.3, 65%)	0.0031 (0.4, 66%)	0.010 (0.6, 69%)
NO_3^-	<0.0001 (0.46, 87%)	<0.0001 (0.50, 78%)	<0.0001 (0.40, 73%)	<0.0001 (0.31, 76%)
Other (by difference) ^a	<0.0001 (2.5, 77%)	0.0028 (1.9, 77%)	<0.0001 (1.8, 67%)	<0.0001 (2.4, 77%)

Note: The first number reported is the p -value resulting from a test of the hypothesis that the mean difference between the log-transformed concentrations measured at Steubenville and at the satellite site under consideration is equal to zero. The mean raw difference between the Steubenville and satellite site concentrations (in $\mu\text{g}/\text{m}^3$) and the percentage of days for which the concentration at Steubenville was greater than that at the satellite site are provided in parentheses. ^aIncludes concentrations of EC and OM.

Table 5. Matrix showing the relationships among components of PM_{2.5} at Steubenville.

	SO ₄ ²⁻	NO ₃ ⁻	NH ₄ ⁺	EC	OM	Other
SO ₄ ²⁻	—	+0.05	+0.71	+0.26	+0.36	+0.34
	—	0.0083	<0.0001	<0.0001	<0.0001	<0.0001
NO ₃ ⁻	+0.05	—	+0.16	+0.03	+0.06	+0.07
	0.0083	—	<0.0001	0.047	0.0061	0.0029
NH ₄ ⁺	+0.71	+0.16	—	+0.17	+0.21	+0.27
	<0.0001	<0.0001	—	<0.0001	<0.0001	<0.0001
EC	+0.26	+0.03	+0.17	—	+0.50	+0.23
	<0.0001	0.047	<0.0001	—	<0.0001	<0.0001
OM	+0.36	+0.06	+0.21	+0.50	—	+0.28
	<0.0001	0.0061	<0.0001	<0.0001	—	<0.0001
Other ^a	+0.34	+0.07	+0.27	+0.23	+0.28	—
	<0.0001	0.0029	<0.0001	<0.0001	<0.0001	—

Note: For each cell, the top number is the R^2 obtained by performing linear regression analysis on adjusted values of the two components, and the sign refers to the slope of the regression line. The bottom number is the p -value resulting from a test of the hypothesis that the slope of the regression line is equal to zero. ^aComputed by difference; does not include EC and OM.

and 0.4, are likely indicative of the effects of regional transport and meteorology on the ambient concentrations of these variables, because it is unlikely that they originate from identical sources. The strong relationship between SO₄²⁻ and NH₄⁺ is expected because particulate SO₄²⁻ is generally collected as NH₄HSO₄ or as (NH₄)₂SO₄. In fact, the summertime molar ratio of NH₄⁺ to SO₄²⁻ ranged from a 25th percentile of 1.8 to a 75th percentile of 2, suggesting that the majority of fine particulate sulfate collected on the filter was fully neutralized. However, this result is potentially due to an artifact by which acidic sulfate collected on the Teflon filter could react with airborne NH₃ during sampling and handling, causing a misrepresentation of the extent to which the actual sampled sulfate was neutralized. Sampling techniques designed for the determination of aerosol acidity have attempted to prevent this potential artifact by employing a denuder to remove NH₃ from the sampled airstream;³⁷ however, the FRM does not use such a denuder. The appreciable correlation observed between EC and OM is possibly attributable to the presence of substantial primary emissions of OM, which likely originates from some of the same sources that emit EC.

Further linear regression analyses were performed to determine whether any component of the PM_{2.5} at Steubenville might be responsible for driving elevated fine particulate concentrations and whether PM_{2.5} or any of its components were related to PM_{10-2.5}. Adjusted values were again used in the analyses. The results are presented in Table 6. Adjusted values of each fine fraction component (x) were regressed against adjusted values of the

Table 6. Regression results for PM_{2.5} components, x , at Steubenville versus PM_{2.5} - x (defined in text) and PM_{10-2.5}.

x	PM _{2.5} - x	PM _{10-2.5}
SO ₄ ²⁻	+0.53 ^a	+0.10
	<0.0001	0.0002
NO ₃ ⁻	+0.08	+0.02
	0.0006	0.098
EC	+0.41	+0.36
	<0.0001	<0.0001
OM	+0.61	+0.36
	<0.0001	<0.0001
PM _{2.5}	—	+0.22
	—	<0.0001

Note: Adjusted values were used. The top number is the R^2 ; the bottom is the p -value resulting from a test of the hypothesis that the slope of the regression line is equal to zero. ^aResults are presented for PM_{2.5}-SO₄²⁻-NH₄⁺.

difference between PM_{2.5} and that fine fraction component (PM_{2.5} - x), because inclusion of the component in the PM_{2.5} mass would inevitably induce a correlation. Additionally, SO₄²⁻ was regressed against PM_{2.5} - SO₄²⁻ - NH₄⁺ (i.e., all mass not attributed to SO₄²⁻ and NH₄⁺) because of the strong relationship between SO₄²⁻ and NH₄⁺. The positive, statistically significant associations observed for SO₄²⁻, EC, and especially OM indicate that elevated concentrations of each of these components tend to be accompanied by elevated concentrations of the combined remaining constituents of PM_{2.5}. Because the associations are observed for all three of these components, it does not appear that any one of them alone is responsible for driving high PM_{2.5} days in Steubenville. The association for NO₃⁻ was not nearly as strong.

The results in Table 6 also reveal that PM_{10-2.5} concentrations are positively and significantly associated with PM_{2.5} concentrations, although PM_{2.5} accounts for only 22% of the variance in PM_{10-2.5} after serial correlations have been removed. Associations between adjusted fine fraction nitrate and sulfate values and adjusted PM_{10-2.5} values are statistically insignificant or relatively weak. However, fine fraction EC and OM are more strongly and significantly correlated with PM_{10-2.5}. This is perhaps attributable to primary source emissions of these carbonaceous species, because coarse PM is predominantly a primary pollutant. Alternatively, it may indicate that carbonaceous particles are distributed over a wide size range in Steubenville.

CONCLUSIONS

Population, industry, and air pollution levels in Steubenville have changed substantially since the landmark Six

Cities Study discovered an association in the 1980s between the city's elevated $PM_{2.5}$ concentrations and adverse health effects. Given the promulgation of a new NAAQS for $PM_{2.5}$ in 1997, data collected from 2000 through 2002 as part of SCAMP provide a much-needed current assessment of $PM_{2.5}$ in the Steubenville region. The characterization of $PM_{2.5}$ and its major ionic and carbonaceous components provided in this paper lays the groundwork for three additional papers in a series that thoroughly analyzes speciated $PM_{2.5}$, gaseous pollutant, and meteorological data collected during SCAMP to enhance our understanding of factors that influence $PM_{2.5}$ in the Steubenville region and to provide insights regarding the design and interpretation of health effects studies.

Data presented here show that although average $PM_{2.5}$ concentrations in Steubenville have decreased by more than $10 \mu\text{g}/\text{m}^3$ since the Six Cities Study, the average $PM_{2.5}$ concentration measured at Steubenville during SCAMP ($18.4 \mu\text{g}/\text{m}^3$) was still $3.4 \mu\text{g}/\text{m}^3$ above the annual $PM_{2.5}$ NAAQS. Hence, a $PM_{2.5}$ reduction strategy will likely be required for the city. Sulfate and organic material are the major constituents of $PM_{2.5}$ in Steubenville, accounting for $\sim 31\%$ and 25% of the total mass, respectively. On a seasonal basis, SO_4^{2-} concentrations track total $PM_{2.5}$ mass concentrations more closely than do OM concentrations; however, regression analyses revealed that both of these components exhibit appreciable positive associations with the remaining constituents of $PM_{2.5}$, suggesting that neither is singly responsible for driving elevated $PM_{2.5}$ concentrations in Steubenville. Spatial analyses indicated statistically significant positive associations for $PM_{2.5}$, SO_4^{2-} , NH_4^+ , NO_3^- , and the collective other components of $PM_{2.5}$ across all pairs of SCAMP monitoring sites (intersite distances ranged from 12 to 80 miles), suggesting the influence of regional meteorology and long-range transport. Local sources in Steubenville on average contributed an estimated $4.6 \mu\text{g}/\text{m}^3$ to the city's ambient $PM_{2.5}$ concentration. Carbonaceous material, crustal material, and trace elements likely account for more than half this local contribution; sulfate accounts for $\sim 1 \mu\text{g}/\text{m}^3$.

Statistical issues such as autocorrelation that have been overlooked by many $PM_{2.5}$ studies were also considered. Analyses presented here show that statistically significant autocorrelation should be considered when examining time series of $PM_{2.5}$ or $PM_{2.5}$ component data collected at frequencies of 1-in-1 to 1-in-4 days, because it can account for a substantial portion of the variability in these time series and lead to incorrect conclusions about statistical significance. ARIMA models generally proved effective in accounting for this autocorrelation, and they were applied in this paper when necessary. Results of

spatial analyses presented here that accounted for autocorrelation were generally consistent with the conclusions of previous descriptive studies. However, the ARIMA examples that were presented, which showed that p -values can change by several orders of magnitude upon properly adjusting for autocorrelation, suggest that analyses attempting to assess statistical significance or precisely quantify relationships among time series of $PM_{2.5}$ data for comparative or predictive purposes must account for autocorrelation to obtain meaningful results.

ACKNOWLEDGMENTS

The authors thank the subcontractors and participating groups who made the 2-yr monitoring program possible. This group includes Optimal Technologies, Ohio University, Wheeling Jesuit University, Saint Vincent College, Franciscan University of Steubenville, the Harvard School of Public Health, and the staff of CONSOL Energy Inc. Research and Development. The outdoor ambient air monitoring portion of SCAMP was funded by the National Energy Technology Laboratory of the U.S. Department of Energy (DOE). Additional funders of SCAMP include the Ohio Coal Development Office within the Ohio Air Quality Development Authority, Electric Power Research Institute, American Petroleum Institute, National Mining Association, American Iron and Steel Institute, Edison Electric Institute, National Institute of Environmental Health Sciences, EPA, and CONSOL Energy Inc. This paper was prepared with the support of DOE Grant DE-FC26-00NT40771. However, any opinions, findings, conclusions, or recommendations expressed herein are those of the authors and do not necessarily reflect the views of the DOE.

REFERENCES

1. Dockery, D.W.; Pope, C.A.; Xu, X.; Spengler, J.D.; Ware, J.H.; Fay, M.E.; Ferris, B.G.; Speizer, F.E. An Association Between Air Pollution and Mortality in Six U.S. Cities; *N. Engl. J. Med.* **1993**, *329*, 1753-1759.
2. Pope, C.A.; Thun, M.J.; Namboodiri, M.M.; Dockery, D.W.; Evans, J.S.; Speizer, F.E.; Heath, C.W. Particulate Air Pollution as a Predictor of Mortality in a Prospective Study of U.S. Adults; *Am. J. Respir. Crit. Care Med.* **1995**, *151*, 669-674.
3. Moolgavkar, S.H. Air Pollution and Daily Deaths and Hospital Admissions in Los Angeles and Cook Counties; In *Revised Analyses of Time-Series Studies of Air Pollution and Health*; Special Report; available on the Health Effects Institute Web site, <http://www.healtheffects.org/pubs-special.htm> (accessed 12/10/03).
4. Dockery, D.W.; Cunningham, J.; Damokosh, A.I.; Neas, L.M.; Spengler, J.D.; Koutrakis, P.; Ware, J.H.; Raizenne, M.; Speizer, F.E. Health Effects of Acid Aerosols on North American Children: Respiratory Symptoms; *Environ. Health Perspect.* **1996**, *104*, 500-505.
5. Delfino, R.J.; Murphy-Moulton, A.M.; Burnett, R.T.; Brook, J.R.; Becklake, M.R. Effects of Air Pollution on Emergency Room Visits for Respiratory Illnesses in Montreal, Quebec; *Am. J. Respir. Crit. Care Med.* **1997**, *155*, 568-576.
6. National Ambient Air Quality Standards For Particulate Matter: Final Rule. *Fed. Regist.* **1997**, *62*, 38651-38760.
7. Spengler, J.D.; Thurston, G.D. Mass and Elemental Composition of Fine and Coarse Particles in Six U.S. Cities; *J. Air Pollut. Control Assoc.* **1983**, *33*, 1162-1171.
8. Laden, F.; Neas, L.M.; Dockery, D.W.; Schwartz, J. Association of Fine Particulate Matter From Different Sources With Daily Mortality in Six U.S. Cities; *Environ. Health Perspect.* **2000**, *108*, 941-947.

9. Schwartz, J.; Dockery, D. Particulate Air Pollution and Daily Mortality in Steubenville, Ohio; *Am. J. Epidemiol.* **1992**, *135*, 12-19.
10. Schwartz, J.; Dockery, D.W.; Neas, L.M. Is Daily Mortality Associated Specifically With Fine Particles?; *J. Air & Waste Manage. Assoc.* **1996**, *46*, 927-939.
11. *Steubenville-Weirton MSA Outlook*; available on the West Virginia University Bureau of Business and Economic Research Web site, http://www.bber.wvu.edu/pdf_files/BBER-2001-11.pdf (accessed 05/03/04).
12. *Ranking Tables for Metropolitan Areas: Population in 2000 and Population Change from 1990 to 2000* (PHC-T-3); available on the U.S. Census Bureau Web site, <http://www.census.gov/population/www/cen2000/phc-t3.html> (accessed 05/03/04).
13. Lipfert, F.W. Trends in Airborne Particulate Matter in the United States; *Appl. Occup. Environ. Hyg.* **1998**, *13*, 370-384.
14. *AirData-Access to Air Pollution Data*; available on the U.S. Environmental Protection Agency Web site, <http://www.epa.gov/air/data/index.html> (accessed 02/26/04).
15. *Quality Assurance Guidance Document 2.12: Monitoring PM_{2.5} in Ambient Air Using Designated Reference or Class I Equivalent Methods*; Human Exposure and Atmospheric Sciences Division: Research Triangle Park, NC, 1998.
16. *Standard Operating Procedure for the Determination of Organic, Elemental, and Total Carbon in Particulate Matter Using a Thermal/Optical-Transmittance Carbon Analyzer*; Research Triangle Institute Environmental & Industrial Sciences Division: Research Triangle Park, NC, 2003.
17. Turpin, B.J.; Lim, H.-J. Species Contributions to PM_{2.5} Mass Concentrations: Revisiting Common Assumptions for Estimating Organic Mass; *Aerosol. Sci. Technol.* **2001**, *35*, 602-610.
18. Jansen, J.J.; Edgerton, E.S.; Hansen, D.A.; Hartsell, B.E. Sampling Artifacts in the Federal Reference Method for PM_{2.5}; Proceedings of the International Conference on Air Quality III, Arlington, VA, September 9-12, 2002.
19. Burton, R.M.; Suh, H.H.; Koutrakis, P. Spatial Variation in Particulate Concentrations Within Metropolitan Philadelphia; *Environ. Sci. Technol.* **1996**, *30*, 400-407.
20. Rizzo, M.J.; Pinto, J.P. Initial Characterization of Fine Particulate Matter (PM_{2.5}) Collected by the National Federal Reference Monitoring Network; Proceedings of the Air and Waste Management Association's 94th Annual Conference and Exhibition, Orlando, FL, June 24-28, 2001.
21. Hansen, D.A.; Edgerton, E.S.; Hartsell, B.E.; Jansen, J.J.; Kandasamy, N.; Hidy, G.M.; Blanchard, C.L. The Southeastern Aerosol Research and Characterization Study: Part 1-Overview; *J. Air & Waste Manage. Assoc.* **2003**, *53*, 1460-1471.
22. Pinto, J.P.; Lefohn, A.S.; Shadwick, D.S. Spatial Variability of PM_{2.5} in Urban Areas in the United States; *J. Air & Waste Manage. Assoc.* **2004**, *54*, 440-449.
23. Granger, C.W.J.; Newbold, P. Spurious Regressions in Econometrics; *J. Econometrics.* **1974**, *2*, 111-120.
24. Milionis, A.E.; Davies, T.D. Regression and Stochastic Models for Air Pollution-I. Review, Comments and Suggestions; *Atmos. Environ.* **1994**, *28*, 2801-2810.
25. Box, G.E.P.; Jenkins, G.M. *Time Series Analysis: Forecasting and Control*; Holden-Day: San Francisco, CA, 1976.
26. *R: ARIMA Modeling of Time Series*; available at <http://rweb.stat.umn.edu/R/library/ts/html/arima.html> (accessed 10/20/03).
27. *2002 Ohio Air Quality Report*; available on the Ohio Environmental Protection Agency Division of Air Pollution Control Web site, <http://www.epa.state.oh.us/dapc/ams/AnaQRO2.pdf> (accessed 10/8/04).
28. Edgerton, E.S.; Jansen, J.J.; Hartsell, B.E. Sources of Carbonaceous Material in PM_{2.5} Based on C¹⁴ and Tracer Analyses; Proceedings of the International Conference on Air Quality IV, Arlington, VA, September 22-24, 2003.
29. Sickles, J.E. A Summary of Airborne Concentrations of Sulfur- and Nitrogen-Containing Pollutants in the Northeastern United States; *J. Air & Waste Manage. Assoc.* **1999**, *49*, 882-893.
30. Faust, B.C. Photochemistry of Clouds, Fogs, and Aerosols. *Environ. Sci. Technol.* **1994**, *28*, 217A-222A.
31. *PM_{2.5} Concentrations and Composition in Pittsburgh: Summary of Ambient Measurements from Pittsburgh Air Quality Study (PAQS)*; available at http://www.netl.doe.gov/coalpower/environment/air_q/docs/Ambient_summary.pdf (accessed 10/20/03).
32. Blanchard, C.L.; Hidy, G.M. Effects of Changes in Sulfate, Ammonia, and Nitric Acid on Particulate Nitrate Concentrations in the Southeastern United States; *J. Air & Waste Manage. Assoc.* **2003**, *53*, 283-290.
33. Anderson, R.R.; Martello, D.V.; Rohar, P.C.; Strazisar, B.R.; Tamilia, J.P.; Waldner, K.; White, C.M.; Modey, W.K.; Mangelson, N.F.; Eatough, D.J. Sources and Composition of PM_{2.5} at the National Energy Technology Laboratory in Pittsburgh During July and August 2000; *Energy & Fuels.* **2002**, *16*, 261-269.
34. Chu, S.-H.; Cox, W.M. Relationship of PM Fine to Ozone and Meteorology; In *Proceedings of A&WMA's 91st Annual Meeting & Exhibition*; A&WMA: Pittsburgh, PA, 1998; Paper 98-RA90A.03.
35. Robinson, A.L.; Khlystov, A.; Wittig, B.; Davidson, C.I.; Pandis, S.N. Characteristics and Sources of PM_{2.5} in the Pittsburgh Region; Proceedings of the International Conference on Air Quality III, Arlington, VA, September 9-12, 2002.
36. Kinney, P.L.; Aggarwal, M.; Northridge, M.E.; Janssen, N.A.H.; Shepard, P. Airborne Concentrations of PM_{2.5} and Diesel Exhaust Particles on Harlem Sidewalks: A Community-Based Pilot Study; *Environ. Health Perspect.* **2000**, *108*, 213-218.
37. Koutrakis, P.; Fasano, A.M.; Slater, J.L.; Spengler, J.D.; McCarthy, J.F.; Leaderer, B.P. Design of a Personal Annular Denuder Sampler to Measure Atmospheric Aerosols and Gases; *Atmos. Environ.* **1989**, *23*, 2767-2773.

About the Authors

Daniel P. Connell is an associate engineer; Jeffrey A. Withum, Ph.D. is a research group leader; Stephen E. Winter is a chemist; and Robert M. Statnick, Ph.D., is a retired director of applied research, all with CONSOL Energy Inc. Research & Development. Richard A. Bilonick, Ph.D., is a senior statistician with CONSOL Energy Inc. Address correspondence to: Daniel P. Connell, CONSOL Energy Inc. Research & Development, 4000 Brownsville Road, South Park, PA 15129; e-mail: danielconnell@consolenergy.com.